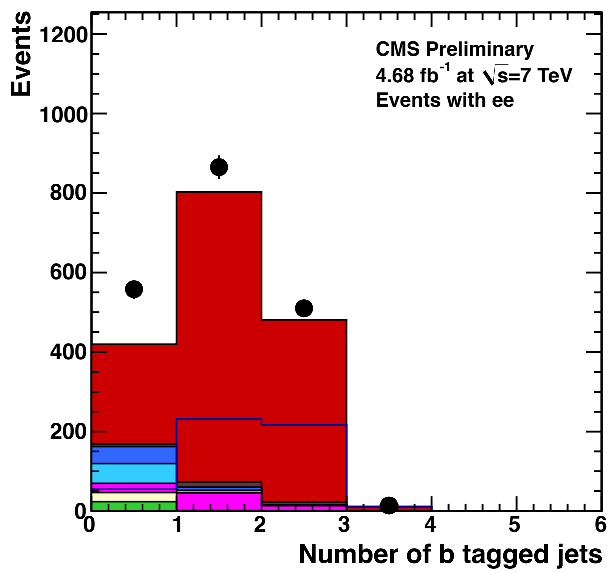
efficiency scale factor for bs and mistag estimate

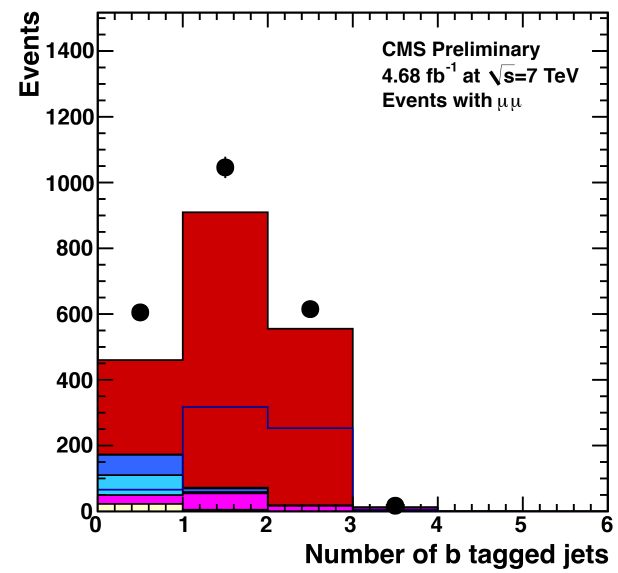
D. Barge, C. Campagnari, P. Kalavase, D. Kovalskyi, V. Krutelyov, J. Ribnik
University of California, Santa Barbara
W. Andrews, G. Cerati, D. Evans, F. Golf, I. Macneill, S. Padhi, Y. Tu, F. Würthwein, A. Yagil, J. Yoo
University of California, San Diego
L. Bauerdick, I. Bloch, K. Burkett, I. Fisk, S. Jindariani, Y. Gao, O. Gutsche, B. Hooberman, J. Linacre
Fermi National Accelerator Laboratory, Batavia, Illinois

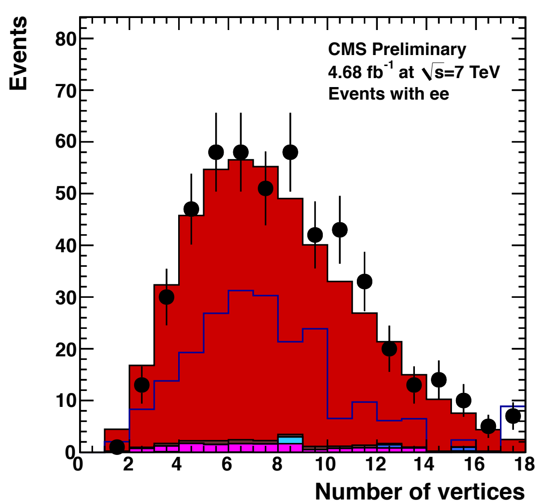
t' meeting
1st December 2011

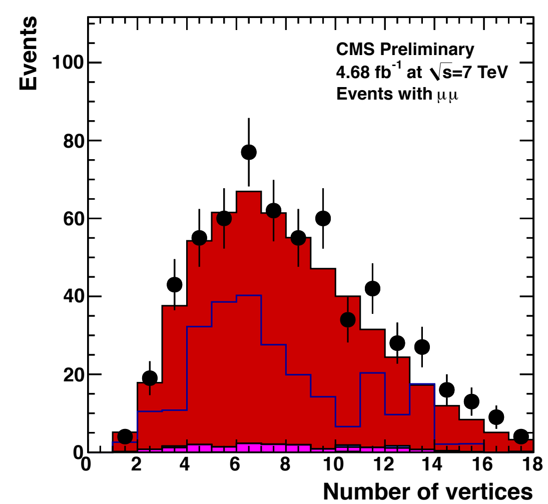
Introduction

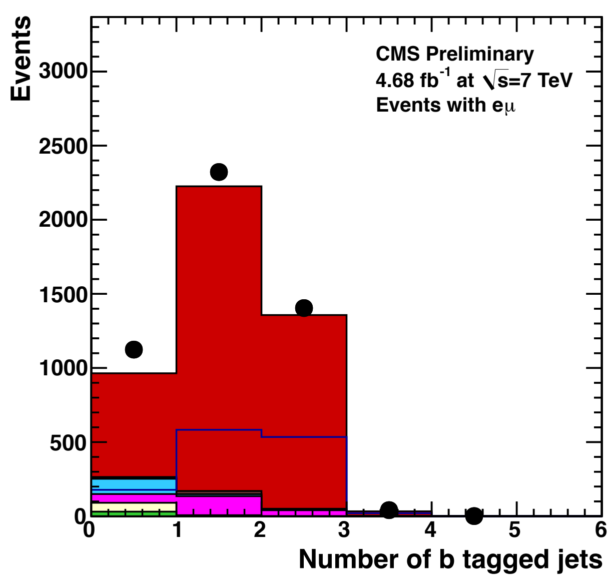
- I showed a plot a month ago with bad data-MC agreement in the 0 and 1 b-tag jet bins. The 2 b-tag bin is our preselection region.

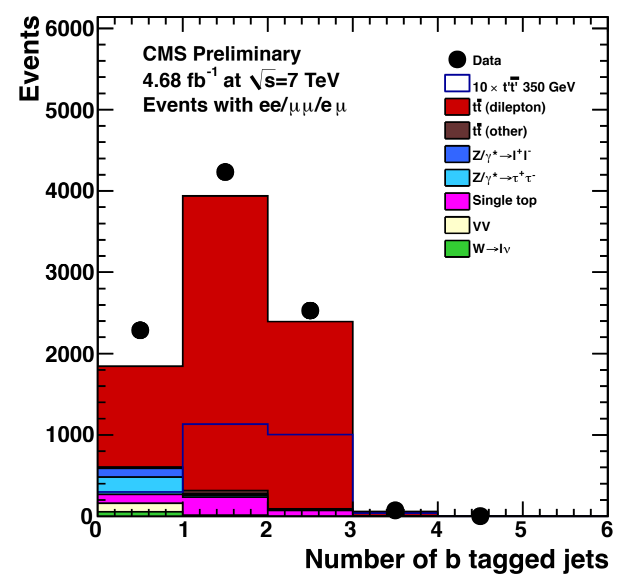


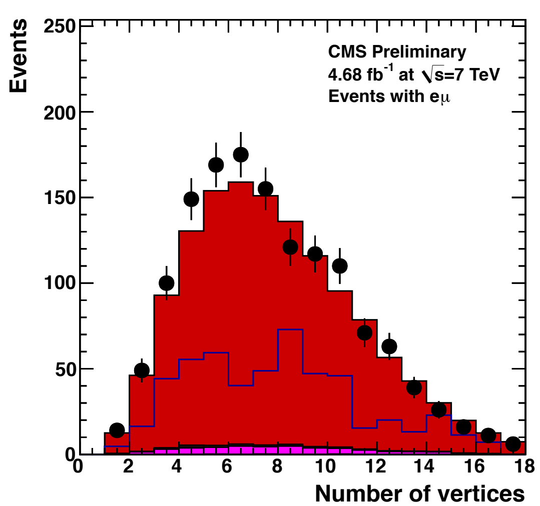


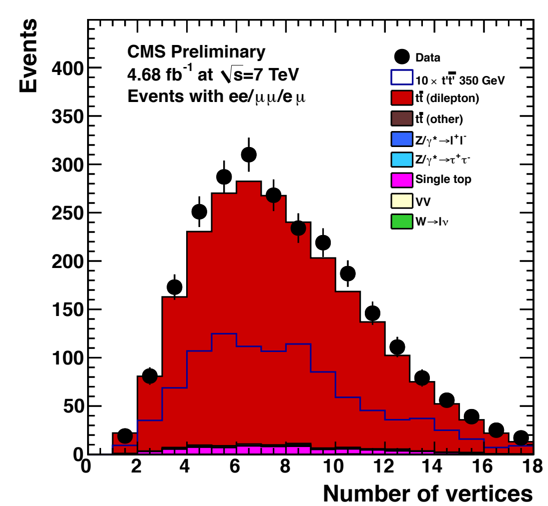












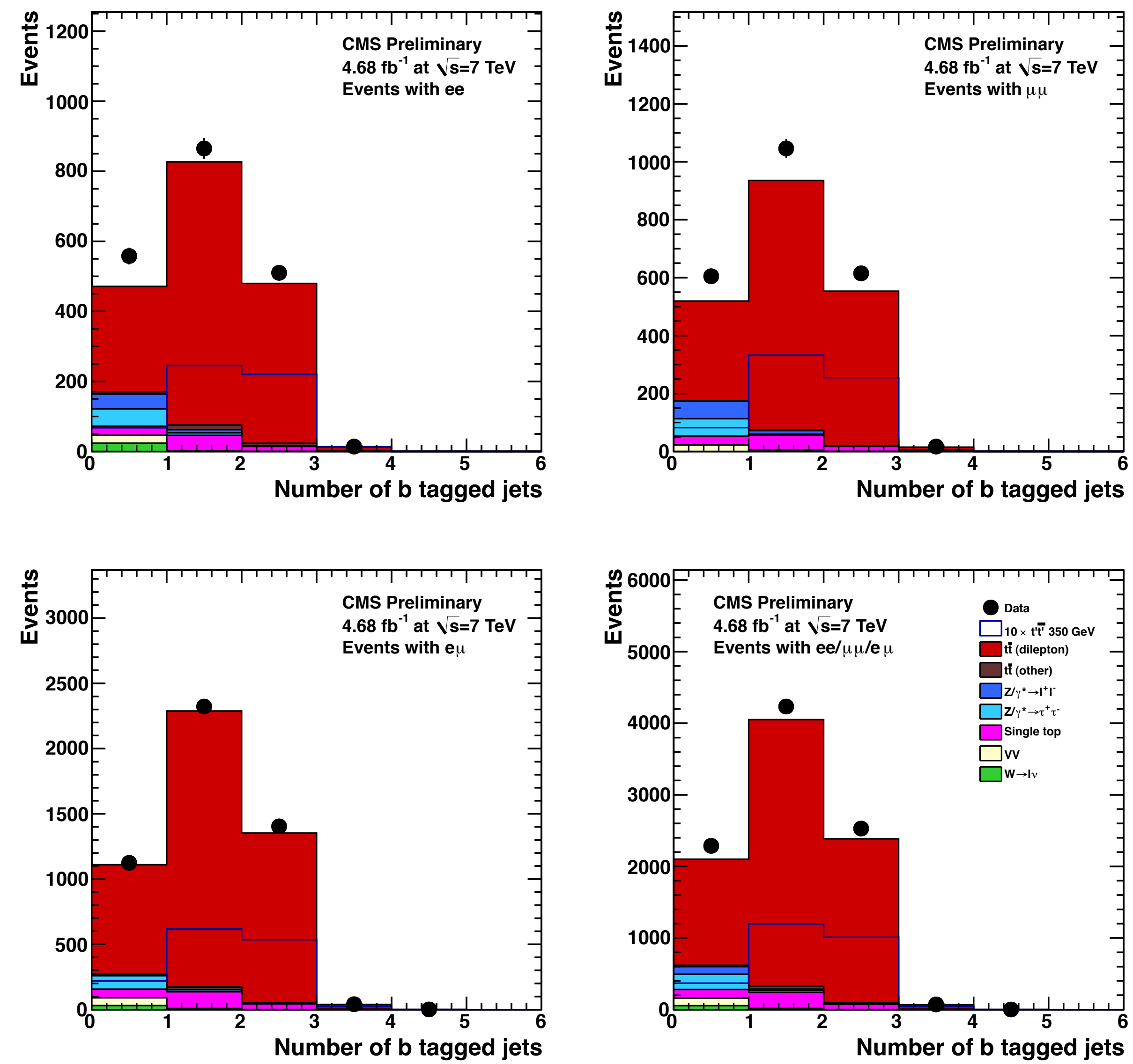
- Plot uses 4.7/fb and the new large Fall11 ttbar sample
  - new vertex weighting: #vertex distribution looks OK (right plots)

data/mc efficiency weighting for bs

- Realised that our btag efficiency weighting is too much of an approximation
  - we use a weight of 0.95 for each btagged jet, where 0.95 is the data/MC scale factor for b-tagging efficiency. This gives a weight of 0.95^2 for events in the signal region.
  - In fact, we should only weight the **real bs** that were tagged.
  - For the remaining b-tagged jets (mistags) the factor of 0.95 is not correct, and we should use the data/MC scale factor for mistags instead (~1.25).
  - Furthermore, we mustn't neglect the real bs that were **not** tagged
    - A data-MC scale factor is also applicable here. For example, an event with 2 real bs but 0 b tags is less likely to appear in MC than in data due to the higher tagging efficiency in MC
  - Finally, there is a similar scale factor for the mistaggable jets that were not mistagged.
- Thus, overall the weight for each event comes from the product of the four scale factors: real b to be tagged * real b not tagged * light jet mistagged * light jet not mistagged
  -
$$\underbrace{F_t^{N_t^b}}_{\text{Num. b-tagged jets matched to real b}} \times \underbrace{F_{nt}^{N_{nt}^b}}_{\text{Num. untagged jets matched to real b}} \times \underbrace{F_m^{N_m^j}}_{\text{Num. b-tagged jets not matched to real b}} \times \underbrace{F_{nm}^{N_{nm}^j}}_{\text{Num. untagged jets not matched to real b}}$$
- These scale factors are taken from http://cms-physics.web.cern.ch/cms-physics/public/BTV-11-001-pas.pdf
- The MC #btags distribution matches the data better when using this weighting

Results

- #btag distributions after weighting
- Note, overall yield with Fall11 madgraph ttbar is 3.3% lower than with Summer11 ttbar
  - not sure why
- Table compares total MC yield to data with the old and new btag efficiency weighting, using either the Fall11 or Summer11 ttbar madgraph sample



- Better agreement with new weighting
- Better agreement with Summer11 ttbar

#btags	Fall11, 0.95^n	Fall11, new	Sum11, 0.95^n	Sum11, new	data
0	1,844.0	2,100.8	1,911.1	2,179.1	2287
1	3,741.4	4,050.6	3,889.6	4,208.7	4233
2	2,393.8	2,385.6	2,453.1	2,445.2	2529
3	54.9	67.3	55.1	67.4	72
4+	2.0	2.8	1.1	1.7	1
All	8,036.1	8,607.1	8,310.0	8,902.2	9122

Fall11 madgraph ttbar sample

- Signal region yields using Fall11 madgraph ttbar sample

Sample	ee	$\mu\mu$	$e\mu$	all
ttprime350	4.9091 ± 0.9196	7.2041 ± 1.0244	10.8035 ± 1.2854	22.9166 ± 1.8834
ttprime400	3.1642 ± 0.4660	4.8099 ± 0.5610	9.9882 ± 0.8301	17.9624 ± 1.1050
ttprime450	2.1583 ± 0.2781	2.3028 ± 0.2698	5.8573 ± 0.4382	10.3185 ± 0.5850
ttprime500	1.3145 ± 0.1526	1.7012 ± 0.1645	2.9942 ± 0.2202	6.0099 ± 0.3143
ttprime550	0.9210 ± 0.1029	0.9156 ± 0.0944	1.8336 ± 0.1373	3.6702 ± 0.1959
ttprime600	0.5142 ± 0.0568	0.5265 ± 0.0544	1.1911 ± 0.0833	2.2318 ± 0.1146
ttdil	0.1659 ± 0.0547	0.1121 ± 0.0446	0.2719 ± 0.0704	0.5499 ± 0.0997
ttotr	0.0137 ± 0.0123	0.0113 ± 0.0113	0.0355 ± 0.0223	0.0605 ± 0.0278
wjets	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
DYee	0.6203 ± 0.6203	0.0000 ± 0.0000	0.0000 ± 0.0000	0.6203 ± 0.6203
DYmm	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
DYtautau	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000	0.0000 ± 0.0000
VV	0.0000 ± 0.0000	0.0085 ± 0.0066	0.4092 ± 0.2894	0.4177 ± 0.2895
tw	0.0000 ± 0.0000	0.1194 ± 0.0849	0.3620 ± 0.1629	0.4814 ± 0.1837
Total MC	0.7999 ± 0.6228	0.2513 ± 0.0968	1.0787 ± 0.3402	2.1299 ± 0.7163
Data	0.0000 ± 0.0000	0.0000 ± 0.0000	1.0000 ± 1.0000	1.0000 ± 1.0000

- Supports earlier results: ttbar no longer dominant background

Updated mistag background estimate

- Prediction of the method compared to true count of events with mistags in MC for each sample

MC Sample	mistag bkg prediction		true # mistag events		pred/true
tt dil (Fall11 madgraph)	0.28	±0.02	0.45	±0.09	61%
tt other (Fall11 madgraph)	0.03	±0.01	0.05	±0.03	51%
w+jets	0.00	±0.01	0		
DYee	0.27	±0.13	0		
DYmm	0.04	±0.04	0		
DYtautau	0.30	±0.16	0		
VV	0.02	±0.01	0.41	±0.29	5%
tw	0.12	±0.02	0.48	±0.18	25%
Sum of MC	**1.04**	**±0.21**	**1.40**	**±0.36**	74%
tt dil (powheg)	0.31	±0.02	0.51	±0.08	61%

- Total prediction covers 74% total mistags, but much less for VV and tW
- Total MC yield (all events): 2.13 ± 0.72 (i.e. 1.40/2.13=66% are from mistags)
  - prediction covers 1.04/2.13 = 49% of total background

82% total Fall11 tt yield is from mistags

- Mistag method background prediction in data: **0.74 ± 0.27**

Closure test in preselection region

- Same as last slide, but for preselection region (without M_{lb}^{min} cut)

Sample	mistag bkg prediction	true # mistag events	pred/true
ttdil	131.62 ±0.33	233.90 ±1.95	0.56 ±0.00
ttotr	1.62 ±0.04	2.69 ±0.21	0.60 ±0.05
wjets	0.094 ±0.101	0.00 ±0.00	
DYee	0.66 ±0.23	0.00 ±0.00	
DYmm	0.65 ±0.30	0.00 ±0.00	
DYtautau	0.94 ±0.26	1.67 ±0.97	0.56 ±0.36
VV	0.24 ±0.04	1.01 ±0.41	0.24 ±0.11
tw	7.05 ±0.15	25.39 ±1.27	0.28 ±0.02
sum of MC	142.87 ±0.60	264.66 ±2.56	0.54 ±0.01
Data	137.70 ±2.71	N/A	0.05 ±0.00

- Prediction covers 54% total mistags
- pred/true consistent with previous slide for the individual samples
- Data prediction matches the sum of MC predictions

- Overall pred/true was 74% in signal region: is this a biased result due to 0 mistag DY events in signal region in MC?

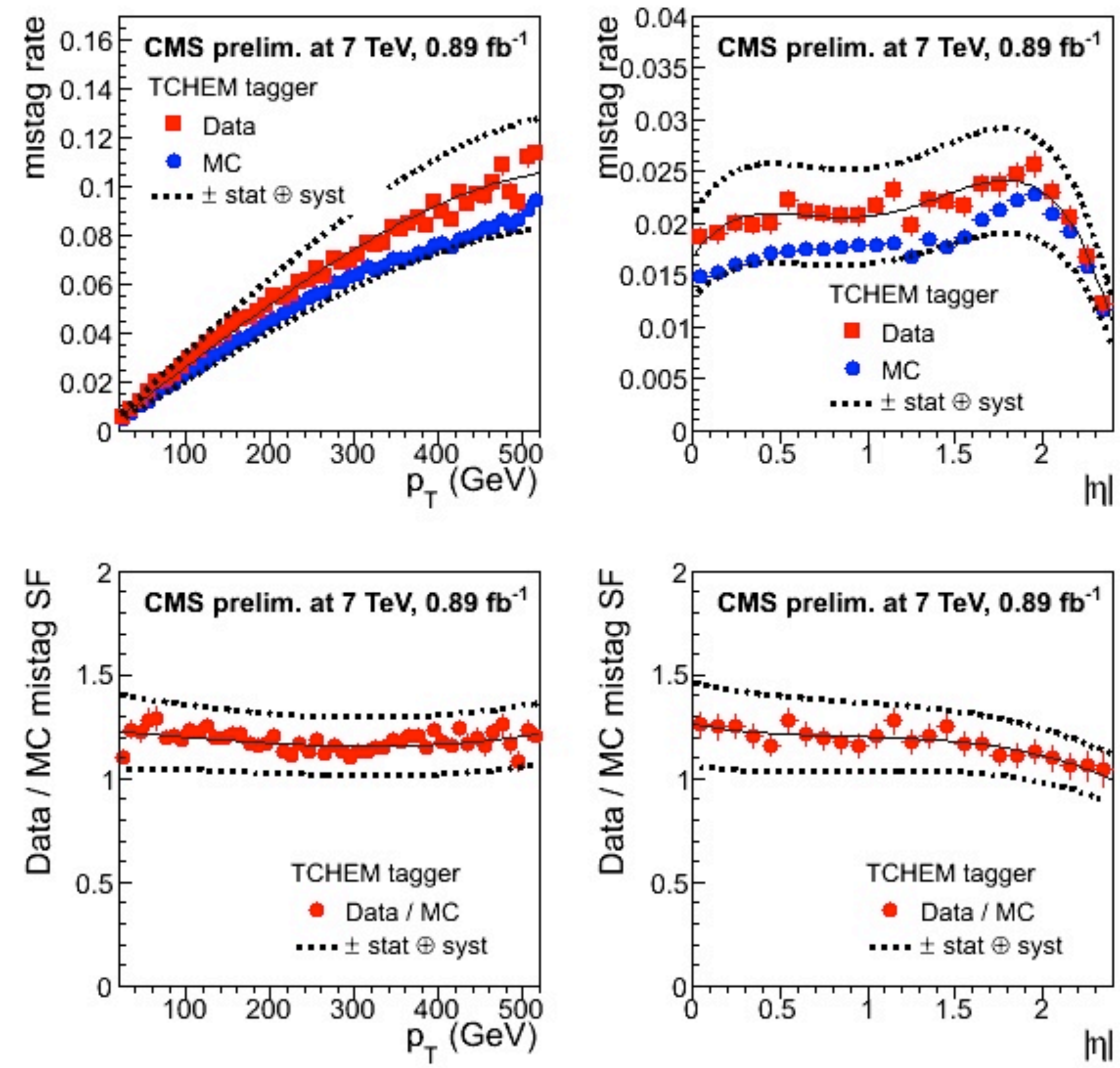
Conclusion

- Method works reasonably well, but final prediction covers 50% (or less) of total background
- We can live with this because the effect of a 50% underestimation of background on our limit is only ~10 GeV
- Assign 100% syst. uncertainty

Backup

Efficiency weighting for bs

- I'm not yet entirely happy with the efficiency weighting
- real b tags are defined as b-tagged jets that match a real b within dR<0.4 (and each real b is only allowed to match 1 btagged jet)
- all other b tagged jets are assumed to be mistags, and weighted as appropriate
  - so they pick up a weight of ~1.25 instead of ~0.95 (from SF plot below)
- Is it OK to call everything else a mistag like this? Some will be badly mismeasured bs.
- To test the effect of the mistag part of the weighting, I set the mistag scale factors to 1 (i.e. only used the 2 real b scale factors).



Results are similar: mistag SFs overall effect small except with ≥3 btags

#btags	b SFs	all SFs
0	2111.1	2100.8
1	4022.6	4050.6
2	2345.0	2385.6
3	57	67.3
4+	2.1	2.8
All	8537.3	8607.1

Simple background estimate idea

- The 1-btag bin data yields now match the MC much better
- Consider using the 1-btag yields to estimate 2 b-tag yields?
  - $N_{predicted}^{2btags} = N_{data}^{1btag} * R_{mc}$
    - $R_{mc} = N_{mc}^{2btags} / N_{mc}^{1btag}$
  - $N_{predicted}^{2btags} = 17 * (2.13/16.69) = 2.17 \pm 0.76$

- But R_{mc} might not be well modelled
- And very poor MC stats ($N_{mc}^{2btags} = 2.13 \pm 0.72$)

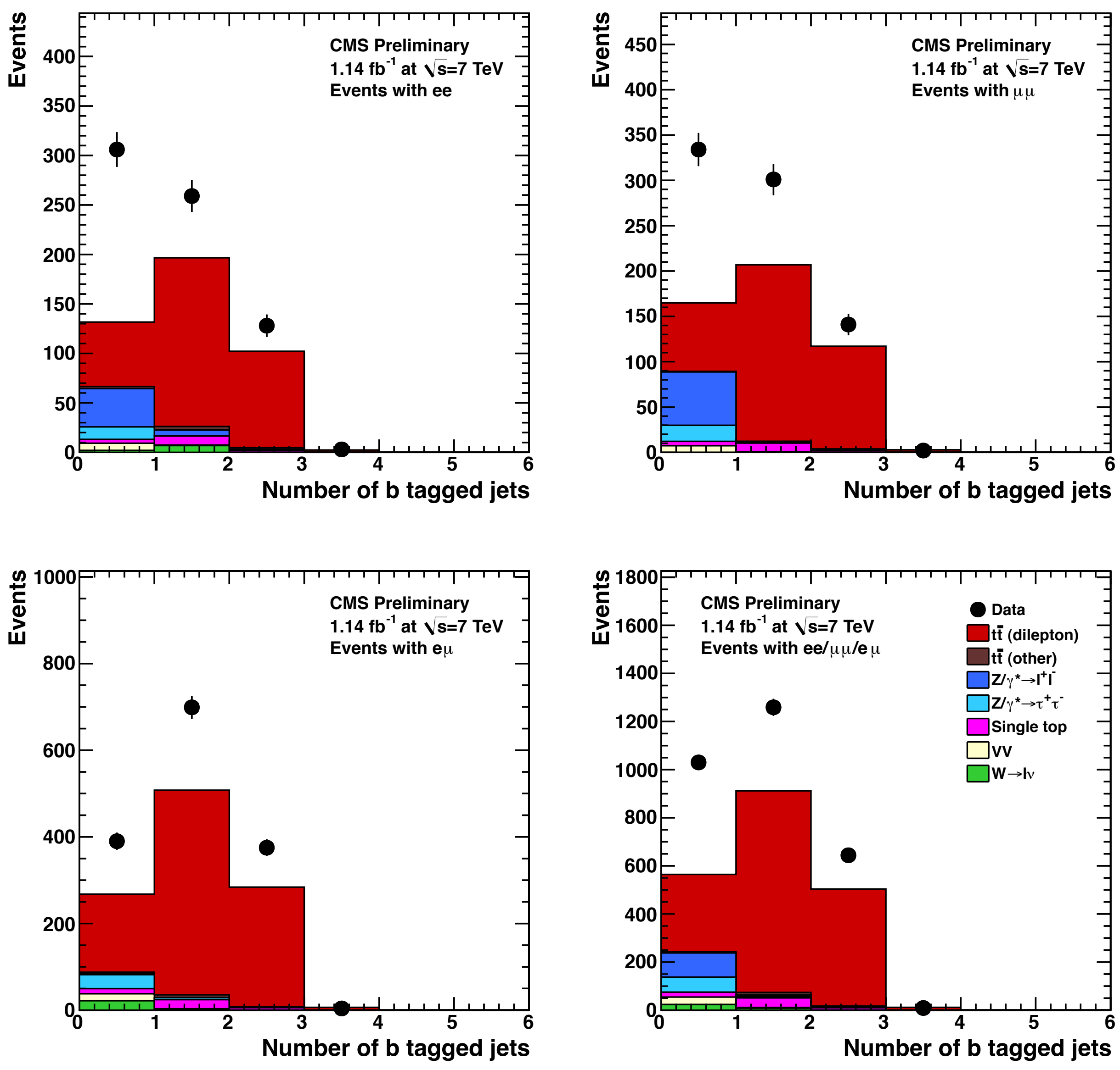
Full closure test results (old)

- Prediction of the method compared to true count of events with mistags in MC for each sample (3.23/fb, Summer11 MC)

MC Sample	mistag bkg prediction		true # mistag events	
tt dil (madgraph)	0.22	±0.06	0.16	±0.16
tt other	0.002	±0.002	0	
w+jets	0.000	±0.004	0	
DYee	0.13	±0.06	0	
DYmm	0.02	±0.02	0	
DYtautau	0.13	±0.06	0	
VV	0.01	±0.01	0.17	±0.12
tw	0.06	±0.01	0.25	±0.10
Sum of MC	**0.57**	**±0.11**	**0.58**	**±0.22**
also looked at powheg pythia tt sample (better stats):				
tt dil (powheg)	0.16	±0.01	0.25	±0.04

- reasonable closure for ttbar and also for sum of MC, but statistics are very limited

- This is the actual plot I showed 1 month ago



- Two reasons it's even worse than the plot on slide 1:
  - weight of 0.95^2 for all events (slide 1 has 0.95^2 for ≥ 2, and 1.00 otherwise)
  - powheg ttbar (overall lower yield than madgraph)